Network Science **8** (1): 42–61, 2020
doi:10.1017/nws.2019.62

CAMBRIDGE
UNIVERSITY PRESS

ORIGINAL ARTICLE

Identifying key players in bipartite networks

Scott W. Duxbury

Department of Sociology, University of North Carolina Chapel Hill, 155 Hamilton Hall, 102 Emerson Drive, Chapel Hill, NC 27514, USA (email: duxbury.5@osu.edu)

Action Editor: Stanley Wasserman

Abstract

Measures of bipartite network structure have recently gained attention from network scholars. However, there is currently no measure for identifying key players in two-mode networks. This article proposes measures for identifying key players in bipartite networks. It focuses on two measures: fragmentation and cohesion centrality. It extends the centrality measures to bipartite networks by considering (1) cohesion and fragmentation centrality within a one-mode projection, (2) cross-modal cohesion and fragmentation centrality, where a node in one mode is influential in the one-mode projection of the other mode, and (3) cohesion and fragmentation centrality across the entire bipartite structure. Empirical examples are provided for the Southern Women's data and on the Ndrangheta mafia data.

Keywords: two-mode network, vulnerability, influence, centrality, disruption

Network scholars have recently shown interest in developing structural measures for bipartite networks (Jasny & Lubell, 2015; Latapy et al., 2008; Opsahl et al., 2010; Opsahl, 2013). First, advances in statistical network models have been extended to affiliation data (Fujimoto et al., 2011; Wang et al., 2009, 2013), opening avenues for inferential analyses of bipartite networks. Second, one-mode projections of bipartite networks (e.g. Breiger, 1974) typically overestimate local structural effects, such as brokerage and clustering (Opsahl et al., 2010; Opsahl, 2013), and sacrifice information contained in the two-mode structure. With these issues and developments comes a need for explicit measures of network structure and node position in bipartite networks.

Notable developments in measurement for bipartite network structure includes global and local clustering coefficients (Opsahl, 2013), geodesic distances (Borgatti & Everett, 1997; Latapy et al., 2008; Newman, 2001), brokerage (Jasny & Lubell, 2015), and determining central cliques (Larsen & Ellersgaard, 2017). However, key player identification has yet to be translated to bipartite networks. Key player identification is of concern for scholars interested in the influence of certain actors over the entire network. For instance, criminologists use key player analysis to locate vulnerabilities in gangs and other organized criminal networks (Duijn et al., 2014; Duxbury & Haynie, 2019). Epidemiologists have applied key player measures to identify actors who are positioned to quickly diffuse a disease or health behavior through a community or population (Chen et al., 2008; Cobb et al., 2010; Valente, 2012), and ecologists have suggested using key player centralities to identify nature reserves that are integral to a functional landscapes (Urban et al., 2009). Extending key player measures to bipartite graphs thus offers a useful scientific tool for research in these fields and related ones.[1]

Borgatti (2006) proposed two measures for key player identification in one-mode networks: fragmentation and cohesion centrality.[2] The former refers to vulnerable positions within a

© Cambridge University Press 2020

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

network, where the latter refers to positions of influence. Unlike brokerage or degree centrality, these measures identify actors who are positioned to yield maximal network disruption when removed (fragmentation centrality)[3] or who are positioned to quickly transmit a signal to the greatest number of actors (cohesion centrality). Such identification strategies are of great interest, for example, when measuring the vulnerability of networks, or when evaluating the potential for a signal to diffuse throughout a network (e.g., disease, new ideas, innovations, or resources).

The goal of this paper is to extend key-player identification to bipartite graphs. It first reviews Borgatti's (2006) measures in depth. It then extends the measures to bipartite networks, extrapolating on the one-mode measure to construct unique metrics for (1) intramodal fragmentation and cohesion centrality, where a node wields influence within a one-mode projection of the bipartite network; (2) cross-modal fragmentation and cohesion centrality, where a node in one mode of the network may be influential in the one-mode projection of the *other* mode in the two-mode network, and (3) total fragmentation and cohesion centrality, where a node in the network is influential across the entire bipartite structure. It uses simulations to illustrate that the proposed measures of fragmentation and cohesion centrality identify distinct sets of nodes compared to alternative measures of influence when working with bipartite networks. An empirical example is provided on the Ndrangheta mafia network for fragmentation centrality and another on Southern Women data (Davis et al., 1941) for cohesion centrality.[4] R code to implement the measures is available at the author's GitHub.

1. The key player problem

Borgatti (2006) proposed fragmentation and cohesion centrality to isolate actors who are either in positions of great vulnerability or in great influence in one-mode networks. The measures focus on the relative distance between actors.

1.1 Fragmentation

Fragmentation centrality calculates disruption potential. The measure is based on the shape of the subgraph that results from a focal actor's removal. Let F be the fragmentation centrality of an actor i in a one-mode network with n actors. The fragmentation centrality F_i is proportional to the average distance between actors when i is removed from the network,

$$F_i = 1 - \frac{2\sum \frac{1}{d_{hj}}}{n(n-1)} \tag{1}$$

where d is the geodesic distance between two actors h and j when a focal actor i is removed from the network. The equation utilizes the reciprocal of d so that $\frac{1}{d_{hj}}$ is equal to 0 when actors are completely disjointed (when d is infinite). The value is subtracted from 1 to ensure that increasing values indicate increasing disruption potential. Like most centrality measures, the value ranges between 0 and 1, where higher values of F_i reflect more vulnerable positions in a network (e.g., their removal results in greater path lengths, on average, between any two randomly chosen actors in the network). By identifying the maximum F_i, a researcher can isolate the actor who is the greatest vulnerability in the network—that is, the actor who, when deleted, yields the greatest possible increase in global mean geodesic distance.

1.2 Cohesion

Like fragmentation centrality, cohesion centrality is also measured by relative distances between two actors. Cohesion centrality measures influence in a one-mode network as indicated by the mean path length between a focal actor and all other actors. Let C be the cohesion centrality of an

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

actor i in one-mode network with n actors. C_i is computed by averaging over the distance between a focal actor and all alters in the network,

$$C_i = \frac{\sum \frac{1}{d_{ij}}}{n} \tag{2}$$

as above, C_i is constrained to range between 0 and 1. Higher values indicate shorter path lengths, on average, from the focal actor to a randomly chosen alter, and therefore greater diffusion potential. By identifying the maximum C_i in a one-mode network, a research can isolate the actor who is positioned to most quickly diffuse a signal throughout the network.

1.3 Key player sets

One of Borgatti's (2006) innovations was to identify key player sets. The utility of these measures is such that they can be updated to identify groups of key players in a network when the influence of a group of actors may be greater than the sum of each actor's influence. This is ideal when diffusion or fragmentation potential are nonlinear—that is, for example, when the removal of the three leading key players may have a weaker effect than the removal of the leading three-player set. The measures can be adjusted by replacing i in each equation with a predetermined set of size k. This results in:

$$F_k = 1 - \frac{2\sum \frac{1}{d_{hj}}}{n(n-1)} \tag{3}$$

for fragmentation centrality, and

$$C_k = \frac{\sum \frac{1}{d_{kj}}}{n} \tag{4}$$

for cohesion centrality. A researcher interested in key sets of players merely determines the size of the desired set. A greedy optimization algorithm can be used to iterate across possible k set combinations and identify the leading k set (Borgatti, 2006; An & Liu, 2016). Since the basic approach to key player set identification does not change in the bipartite context, this issue will not be discussed at length. However, the measurements proposed for identifying key players in two-mode networks can easily be adapted for key player sets in the same fashion as those above.

2. Key players in bipartite networks

A core component of key player identification is the relative distance of a focal actor to alters. Yet, there is little consensus on the best way to measure distance in a two-mode network. To derive key player metrics in the bipartite context, the paper proceeds by first discussing how to measure distance in a bipartite network before developing measurements for fragmentation and cohesion centrality.

2.1 Distance

The conventional strategy for assessing distance in a bipartite network is to first project each mode of the network and then compute distance between actors in each simple one-mode projection (e.g. Newman, 2001, 2010). This approach offers the benefit of parsimony and allows researchers to consider the implicit relationships between actors who belong to the same mode of the network (Breiger, 1974). However, one-mode projections often overestimate some structural features,

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

such as local clustering (Opsahl, 2013). More pertinent to the detection of key players in bipartite networks is that simple one-mode projections sacrifice information contained in two-mode networks. Such information loss has obvious implications for key player identification, where over or underestimating distance in the one-mode projections may affect which actors are identified as influential. Given a two-mode network which can be represented by an affiliation matrix with m actors in the first mode and p actors in the second, one possible approach is to reconfigure the $m \times p$ affiliation matrix into a square $(m + p) \times (m + p)$ matrix, such that both modes appear in both the columns and the rows of the reconfigured matrix (Borgatti & Everett, 1997). However, this approach yields almost identical results as the one-mode projection strategy.

An alternative strategy to account for distances in projections of bipartite networks is to weight the edges in the projected networks by mutual affiliations in the bipartite network (Newman, 2001; Larsen & Ellersgaard, 2017). Newman (2001) defines this measure as:

$$w_{ij} = \sum_{q} \frac{1}{n_q - 1} \tag{5}$$

where, given a projection of ij edges, the weight w for an ij edge is proportional to the degree centrality of q (n_q), who is connected to both i and j in the unprojected two-mode network. Since an ij edge can only exist if $n_q \geq 2$, there is no risk of dividing by infinite values. The distance between two actors can then be computed by summing over the weights of each edge on the shortest path between i and j in the weighted one-mode projection.

This approach is appealing for measuring distance in bipartite networks because the weighting procedure preserves information from the two-mode network. In some instances, however, the strategy of summing over weighted edges to calculate distance raises the possibility that smallest value or "shortest" path between two actors may not actually be the path with the smallest number of intermediaries. Considering this issue in the case of weighted one-mode networks, Opsahl et al. (2010) propose a solution to considering both edge-weights and the number of intermediaries between actors in a weighted one-mode network. Here, the researcher specifies a tuning parameter which assigns greater or lesser priority to edge weights. When applied to a weighted one-mode projection, the shortest path length d between i and j can be expressed as

$$d^b_{ij} = \frac{1}{w^a_{iz}} + \cdots + \frac{1}{w^a_{zj}} \tag{6}$$

where superscript b distinguishes the measure of distance in bipartite structures from simple distance d, z represents the intermediary nodes on the path between i and j, and α is a tuning parameter. Values of α larger than 1 place greater importance on the weights in the projected network, while values below 1 assign greater weight to the number of intermediaries in a path. Since a case can be made that both the strength of the weighted ties and the number of intermediaries in a one-mode projection are important when evaluating distance in a bipartite network, it is often appropriate to assign a value of 1 to α. When a value of 1 is used, the distance measure reduces to the summation approach discussed by Newman (2001), and can be represented as:

$$d^b_{ij} = \frac{1}{w_{iz}} + \cdots + \frac{1}{w_{zj}} \tag{7}$$

Equation (7) will be used for all empirical analyses in this study. Opsahl et al. (2010) and Agneessens et al. (2017) provide guidance on selecting a substantively meaningful value for α.

To summarize, distance can be measured in bipartite networks by first weighting the edges in one-mode projections by the number of mutual affiliations in the bipartite network and then summing over the edge weights for all intermediary links in the path between two actors. The result is a distance measure which accounts for proximity in bipartite network structure as well as the absolute path lengths between actors in the one-mode projections.

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

2.2 Fragmentation

In a two-mode network, the projection between actors in the first mode, the projection between actors in the second mode, and the affiliations between modes are distinct relationships. Separate key player measures should be created to summarize the relationships in both projections and the overall bipartite structure (Latapy et al., 2008). There are three ways in which actors' influence can be conceptualized in a bipartite network. First, actors may be influential within the one-mode projection to which they belong (*intramodal* influence). Second, actors may be influential in the one-mode projection to which they *do not* belong (*cross-modal* influence). Third, actors may be influential in the entire bipartite network structure—that is, across both projected networks (*total* influence).

2.2.1 Intramodal fragmentation

Intramodal fragmentation centrality considers the vulnerability of actors within a one-mode projection of a bipartite network. A simple way to extend fragmentation centrality in one-mode networks to intramodal fragmentation centrality is to replace the simple distance measure in Equation (1) with the distances obtained from Equation (7). In this approach, an actor is deleted in one mode of the network. Next, the network is projected into a weighted one-mode network, with edge weights assigned using Equation (7). The researcher then computes Equation (1) and repeats the procedure for all actors in the network.

While intuitively appealing, this approach raises some interpretive issues. Namely, $\frac{2\sum \frac{1}{d^b_{hj}}}{n(n-1)}$ will not be constrained to achieve a maximum value of 1. As a consequence, subtracting the value from 1 will yield negative values in some circumstances (when $\frac{2\sum \frac{1}{d^b_{hj}}}{n(n-1)} > 1$), and positive values in others (when $\frac{2\sum \frac{1}{d^b_{hj}}}{n(n-1)} < 1$). In applied research, this may sometimes yield negative intramodal fragmentation centralities for actors who are in vulnerable locations, and positive intramodal fragmentation centralities for actors who are not vulnerable.

One way to resolve this issue is to construct intramodal fragmentation centrality by taking the reciprocal of $\frac{2\sum \frac{1}{d^b_{hj}}}{n(n-1)}$. This provides an alternate approach to ensuring that increasing values reflect greater fragmentation potential. While values will exceed 1, they will not dip below 0. Let F^b_i be the intramodal fragmentation centrality for an actor in a one-mode projection of a bipartite network. F^b_i can be defined as,

$$F^b_i = \frac{n^b\left(n^b-1\right)}{2\sum \frac{1}{d^b_{hj}}} \tag{8}$$

where n^b is the number of actors in the intramodal projection and the numerator of the equation indicates the number of possible ties in the projected network. As in the case of Equation (1), 1 is divided by d^b_{hj} to avoid dividing infinite values, and to ensure that disconnected components of the network yield $\frac{1}{d^b_{hj}} = \frac{1}{Inf} = 0$. Thus, infinite values of F^b_i are only obtainable in completely disconnected graphs where all nodes are isolates. Higher values reflect greater disruption potential, where removing an actor with the largest F^b_i will yield the greatest obtainable increase in distances in the projected network to which i belongs. The benefit of this approach is that it standardizes the distance measure by the size of the network. Thus, intramodal fragmentation centralities in networks of different sizes can be compared. The actor with the greatest intramodal vulnerability can be identified by maximizing over F^b_i.

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

2.2.2 Cross-modal fragmentation

An interesting property of two-mode networks is that removing an actor in one mode of the network will have a ripple effect on the projection of the alternate mode. Thus, fragmentation measures can be constructed to evaluate how node deletion results in cross-modal disruption. Following the approach above, cross-modal fragmentation centrality can be computed by first deleting a node in one mode of the bipartite network. The alternate mode of the network is then projected and weights computed. The distances between actors in the projected network are summed as $\sum \frac{1}{d_{hj}^c}$, where superscript c denotes that distances are being computed in the cross-modal projection—that is, the projected mode of the network of which i is *not* a member. The score is then standardized by the number of actors in the cross-modal projection. Let F_i^c indicate the cross-modal fragmentation centrality of an actor i,

$$F_i^c = \frac{n^c\,(n^c - 1)}{2\sum \frac{1}{d_{hj}^c}} \tag{9}$$

where d_{hj}^c is the distance score for ties connecting alternate mode actors h and j in the cross-modal projection, and n^c is the number of actors in the cross-modal projection. The interpretation of this measure is that removing an actor with higher cross-modal fragmentation centrality will yield larger increases on the distance between actors in the cross-modal projection. The actor with the greatest cross-modal vulnerability can be identified by maximizing over F_i^c.

2.2.3 Total fragmentation

Latapy et al. (2008) note that there is not a good measure of average distance for the two-mode structure of a bipartite network; only measures for one-mode projections. Consequently, creating a key player statistic for the affiliation structure poses a problem. A possible means to translate the key player statistic to bipartite structure is to create an aggregate measure of fragmentation based on the affiliations between first and second mode actors. Such a measure would estimate how an actor i disrupts the $p \times p$ projection as well as the $m \times m$ projection. The strategy of projecting both matrices and combining their measures is similar in concept (though not process) to the dual projection approach of Everett & Borgatti (2013); see also Everett (2016).

The concern with this measure is to gauge the effect an actor's removal has on both modes of the network. Let F_i^* be the total fragmentation centrality for an actor i. Since F_i^b and F_i^c are standardized by both the size of the projected networks and the distance between nodes in the affiliation structure, these measures can simply be averaged to create a composite fragmentation centrality statistic. The resulting value indicates the average disruption potential of an actor across both projections,

$$F_i^* = \frac{F_i^b + F_i^c}{2} \tag{10}$$

where higher values reflect greater average disruption across both projected networks. The actor with the greatest average vulnerability across both modes of the bipartite network can be identified by maximizing over F_i^*.

2.3 Cohesion

2.3.1 Intramodal cohesion

As in the case of intramodal fragmentation centrality, the simplest way to calculate intramodal cohesion centrality is to replace d_{ij} in Equation (1) with d_{ij}^b (Equation (7)). Here, the intramodal cohesion centrality of an actor is the sum of the distances in a projected network divided by the

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

number of actors in the projection. Increasing values indicate greater cohesion, such that actors with greater cohesion centrality are positioned to diffuse a signal in the shortest amount of time. As above, the value is not constrained to sum 1.

$$C_i^b = \frac{\sum \frac{1}{d_{ij}^b}}{n^b} \tag{11}$$

Like fragmentation centrality, the cross-modal cohesion score for actors can also be constructed. Cross-modal cohesion represents the reachability of an actor i to actors in the cross-modal projection. As the actors' cross-modal cohesion centrality increases, their ability to quickly transmit a signal to actors in the cross-modal projection also increases.

2.3.2 Cross-modal cohesion

The approach to calculating cross-modal cohesion is somewhat different from cross-modal fragmentation. For cross-modal cohesion, the concern is to summarize the reachability of an actor i to all actors in the cross-modal projection. One way to do so would be to calculate cross-modal cohesion as the average intramodal cohesion centrality of alters in the cross-modal projection. The logic of this strategy is that the actor who is most able to diffuse a signal to alters in the cross-modal projection will be the actor who is highly connected to alters who are influential within their own respective intramodal projection. In this approach, the intramodal cohesion of alters is standardized by the number of alters nominated. Let C_i^c be the cross-modal cohesion centrality of i,

$$C_i^c = \frac{\sum C_j^b}{n_j} \tag{12}$$

where the numerator is the summation of the intramodal cohesion centralities for all of i's alters in the alternate mode of the two-mode network. The denominator is the number of alters in the numerator. The benefit of this measure is that it summarizes the diffusion potential of a focal actor's alters. The problem, however, is that actors with more connections will often be penalized for their connectivity because a large number of lower-scoring alters will offset the influence of a handful of high-scoring alters. A simple way to fix this issue is to standardize the measure by the total number of actors in the cross-modal projection. Doing so ensures that highly connected actors are not penalized for their connectivity. This offers the final measure of cross-modal cohesion centrality,

$$C_i^c = \frac{\sum C_j^b}{n^c} \tag{13}$$

where the denominator represents the total number of actors in the cross-modal projection. Maximizing the measure identifies the actor who is most able to diffuse a signal in the cross-modal projection.

2.3.3 Total cohesion

As in the case of fragmentation centrality, a composite measure for two-mode network cohesion can be constructed. The question is how to identify nodes that are influential in the two-mode structure and across both projections. Let C_i^* be the total cohesion centrality for the two-mode structure. Since C_i^b and C_i^c are both standardized by the number of actors in the respective one-mode projections, C_i^* can be computed by averaging the cross-modal and intramodal cohesion centrality of i. The resulting value measures both an actor's influence in its one-mode projection and the influence of the actor's alters in the cross-modal projection. Higher values indicate greater total cohesion centrality:

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

$$C_i^* = \frac{C_i^b + C_i^c}{2} \tag{14}$$

Maximizing the measure identifies the actor who is able to most quickly diffuse a signal across both projections of the two-mode network.

2.4 Key player sets

The proposed measures can also be used to identify key player sets with little adjustment. The approach would mirror that described in the first section, where a set of size k is assigned and the measures are computed for each unique combination of k actors. Of course, this will be computationally intensive if the researcher is interested in F^* or C^* or the network is large.

3. Proof of concept

A simple way to demonstrate the utility of the measure is to compare the proposed measures to simple calculations of key player statistics which do not account for distance in two-mode networks. To do so, I first simulate a random bipartite network with density varying randomly between 0.05 and 0.35. Each network has 65 total actors, 50 in the first mode and 15 in the second. I then calculate the fragmentation and cohesion centrality for each network, using both the equations proposed above as well as "simple" projections which do not adjust for distance in bipartite networks. In the simple projections, cross-modal fragmentation and centrality statistics are calculated using Equations (9) and (13), and intramodal key player statistics are calculated using Equations (1) and (2). The total fragmentation and cohesion centralities are computed using Equations (14) and (10).

Each simulation identified the actors with the largest intramodal, cross-modal, and total fragmentation and cohesion centralities. Identifiers for these actors were then stored for analysis. I computed simple matching coefficients between the vectors of identified actors to assess how frequently the bipartite and one-mode measures identify the same actors. To contextualize these matches against random chance, I also compute correlation coefficients between the vectors of identified actors for each strategy. With 500 iterations, the results reported in Table 1 are the matching coefficients and correlations for key player identification in 500 random bipartite networks. In the analyses below, subscript i denotes actors in the first mode of the network with 50 nodes, while subscript q denotes actors in the second mode of the network with 15 nodes.

Beginning with fragmentation centrality, the two measures are most consistent for intramodal fragmentation centrality in the first mode with 50 actors. Here, both approaches identified the same actors in 15.8% of cases and are correlated at 0.319 ($p < 0.001$). The measures are less consistent when calculating intramodal fragmentation centrality in the second mode with 15 actors, where the two approaches only identified the same actors in 5% of cases ($r = -0.052, p < 0.001$). Turning to cross-modal fragmentation, the proposed measure and simple one-mode projection almost never identified the same actors. F_i^c had a 0% similarity between the simple measures and the proposed approach ($r = -0.024, p > 0.1$). Similarly, there was only 6% similarity for $F_q^c (r = -0.05, p > 0.1)$. Aggregating the intramodal and cross-modal measures to F^*, the one-mode projection and proposed approach are correlated at 0.126 ($p < 0.001$) with 16.4% similarity. This indicates that while there is some overlap between the proposed approach and the one-mode projections for fragmentation centrality, the two measures typically differ.

Results for cohesion centrality are less divergent but tell a similar story. One-mode projections and the proposed method for evaluating intramodal cohesion centrality in the first mode identify the same actors in 28.2% of cases ($r = 0.292, p < 0.001$). However, much like second mode intramodal fragmentation, the % similarity intramodal cohesion centrality in the second

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

Table 1. Comparison of key player measures

	% similarity	Correlation
Intramodal		
Large mode (F_i^b)	15.8	0.319†
Small mode (F_q^b)	5.00	−0.052
Large mode (C_i^b)	28.2	0.292†
Small mode (C_q^b)	0.200	0.069
Cross-modal		
Large mode (F_i^c)	0.000	−0.024
Small mode (F_q^c)	0.600	−0.050
Large mode (C_i^c)	71.8	0.743†
Small mode (C_q^c)	84.6	0.815†
Total		
Fragmentation (F^*)	16.4	0.126†
Cohesion (C^*)	60.4	0.700†

*, $^\dagger p < 0.001$. % similarity is a simple matching coefficient multiplied by 100

mode is very low (0.2) and the two approaches are correlated at 0.069 ($p > 0.1$). Simple projections perform best when identifying cross-modal cohesion scores. Here, the % similarity for cross-modal cohesion centrality in the first mode is 71.8, and the two approaches are correlated at 0.743 ($p < 0.001$). Similarly, when considering cross-modal cohesion centrality in the second mode, the two measures identify the same actor in 84.6% of simulations ($r = 0.815, p < 0.001$). Aggregating the intramodal and cross-modal measures, total cohesion centrality is correlated at 0.7 ($p < 0.001$) for the two approaches with 60.4% similarity. Taken in sum, these results demonstrate that the proposed measures routinely identify distinct sets of key players. While the two strategies show greater overlap when considering cohesion centrality, there is still substantial variation between the two methods.

Next, I conducted a second simulation varying the properties of the simulated bipartite networks across a range of global characteristics. The networks were simulated with random density between 0.01 and 0.3 and with a random number of nodes in each mode (between 10 and 100). I then computed the simple and weighted fragmentation and cohesion centralities for each network. I also computed the degree centrality for the entire network and the betweenness and degree centrality for each projection.[5] With 500 replications each, the second simulation reports the results from multiple measures of influence across 500 simulated bipartite networks. To assess the performance of the proposed measures, I recorded the number of cases where the actor with maximum bipartite key player centrality also maximized an alternative centrality measure. Unmatched cases indicate that simpler measures identified different actors than the bipartite key player statistics.

Figure 1 provides conditional density plots comparing the weighted fragmentation centralities to betweenness centrality, degree centrality, and unweighted fragmentation centrality. For intramodal fragmentation centrality, there were no matches in 77% of networks, meaning that the intramodal fragmentation centrality uniquely identified actors in four out of five simulated networks. For total fragmentation centrality, the maximum total fragmentation centrality was not identified by any other measure in 86% of cases, and cross-modal fragmentation centrality uniquely identified leading actors in 95% of cases. The panels in Figure 1 provide insight into the network properties where the differences between intramodal fragmentation centrality and alternative measures are greatest. Across the panels, the probability of identifying the same leading actor increases as networks increase in either size or density. The relationship between the difference in mode size and the probability of identifying the same actor with another measure is

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

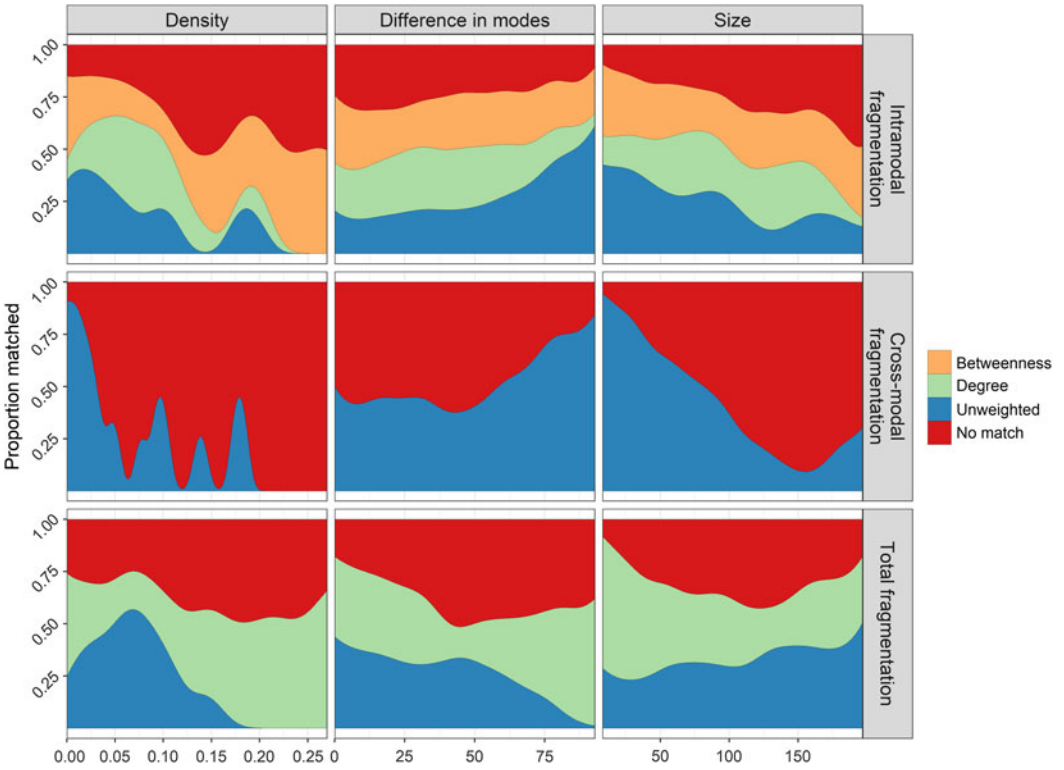


Figure 1. Conditional density plots comparing fragmentation centrality to alternative measures of influence. Y axis is the proportion of networks where the actor with the maximum weighted fragmentation centrality was equivalent to an alternative measure of influence. The difference in mode size is the absolute difference. The unweighted measure is the simple fragmentation centrality.

negative, indicating that the measures become more balanced as the absolute difference in mode size decreases.

As in the prior simulation, there is greater overlap between the cohesion centrality measures and simple measures of influence. Although the maximum intramodal cohesion centrality identifies unique actors in 95% of networks, the maximum cross-modal cohesion centrality only identifies unique actors in 23% of networks, and the maximum total cohesion centrality identifies unique actors in 26% of networks. Figure 2 plots the conditional densities for each measure. Trends in Figure 2 for intramodal cohesion centrality are similar to Figure 1: As network density and network size increase, so does the probability that alternative measures of influence will be unable to identify the same leading actor as the intramodal cohesion centrality. However, there is no clear relationship between network properties and cross-modal cohesion centrality. Total cohesion centrality tends to uniquely identify leading actors in sparse networks, and only rarely identifies unique leading actors in dense networks. There is no clear relationship between differences in mode size and cohesion centrality across measures. These results suggest that the researchers may be able to identify actors who are able to quickly diffuse a signal across the bipartite structure (total cohesion centrality) using simple degree centrality, and that cross-modal cohesion centrality can often be approximated with simple cohesion centrality. That said, alternative measures of influence yielded different results than weighted cross-modal and total cohesion centrality measures in roughly one out of every four simulated networks, which is a non-negligible margin.

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

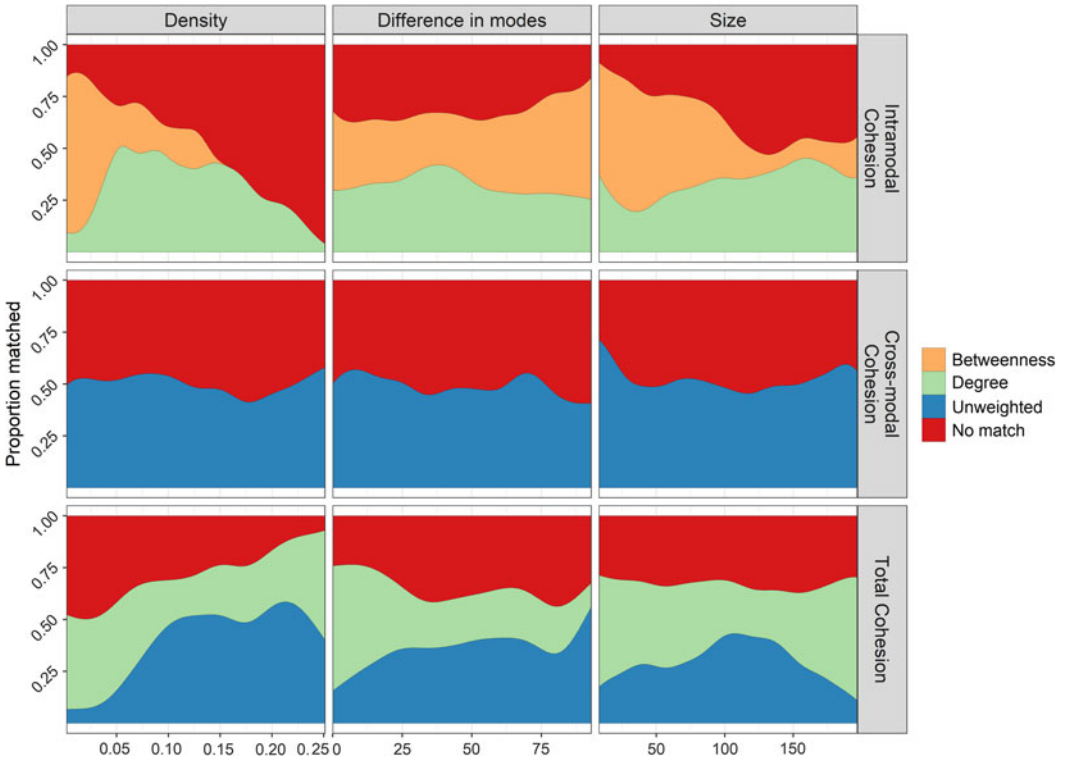


Figure 2. Conditional density plots comparing cohesion centrality to alternative measures of influence. Y axis is the proportion of networks where the actor with the maximum weighted cohesion centrality was equivalent to an alternative measure of influence. The difference in mode size is the absolute difference. The unweighted measure is the simple cohesion centrality.

Figure 3 correlates the weighted key player measures with the unweighted measures. Panel A reveals that there is little association between the weighted and unweighted fragmentation centrality measures, regardless of network structure. Consistent with the results for leading actor identification, Panel B shows that weighted and unweighted intramodal cohesion centrality tend to be weakly-to-moderately correlated associated. However, both total and cross-modal cohesion centrality measures are highly correlated with their unweighted counterparts. These results suggest that, if a researcher is simply interested in aggregating network statistics and is not interested in leading actor identification, unweighted cross-modal and total cohesion centrality may be sufficient.

In sum, results illustrate that fragmentation centrality and intramodal cohesion centrality overwhelmingly identify unique actors when compared to alternative measures, while degree centrality and unweighted cohesion centrality identify similar actors to cross-modal and total cohesion centrality in many cases. Since cross-modal cohesion centrality is an average measure of alters' intramodal cohesion centrality, the comparability of weighted and unweighted cross-modal cohesion centralities indicates that the ranking of average intramodal cohesion centralities across groups of nodes tend to be similar for the weighted and unweighted measures. This similarity, in turn, drives the similarity between the weighted and unweighted total cohesion centralities. Indeed, this is reflected in the high correlations between unweighted and weighted cross-modal and total cohesion centralities. Results also suggest that the divergence between the weighted key player measures metrics increases as network density and size increase for most key player measures, but declines as the absolute difference in mode size increases. Researchers should

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

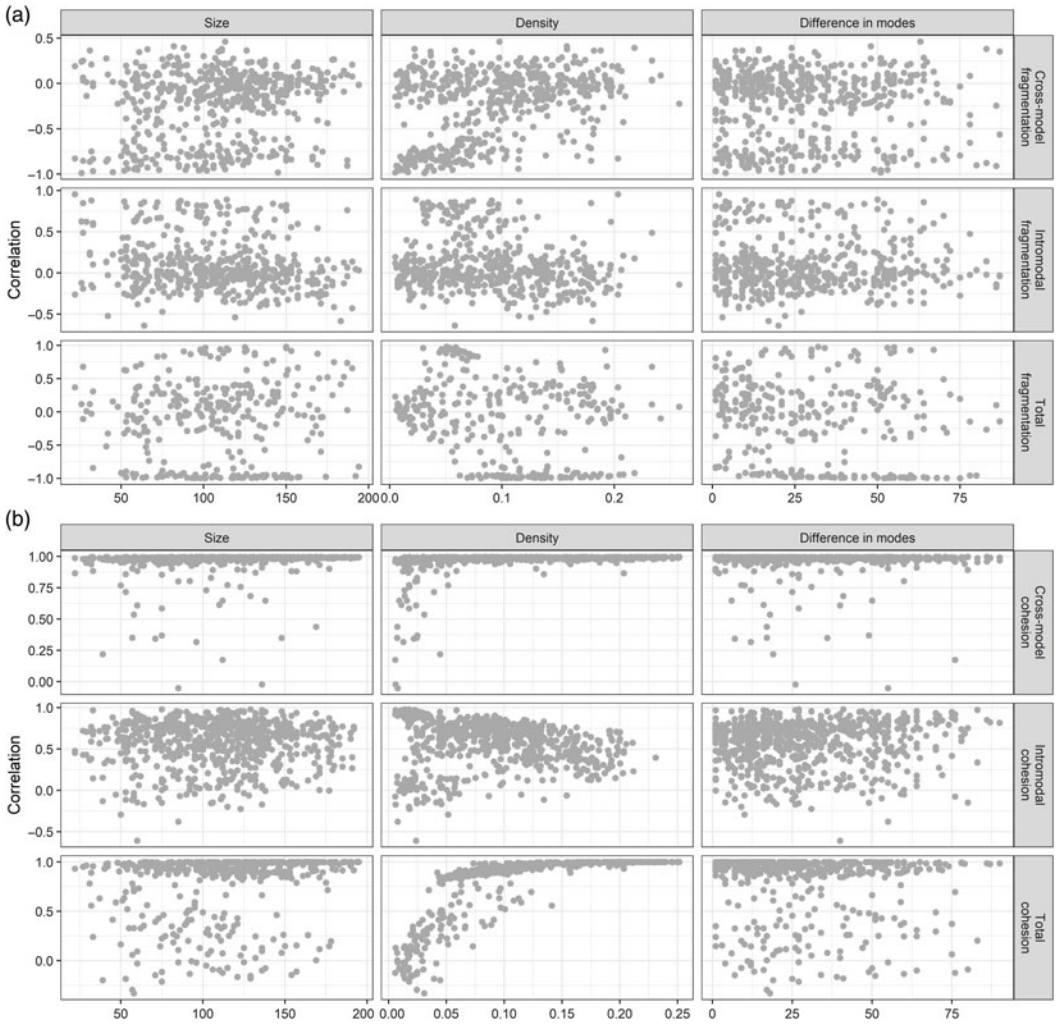


Figure 3. Correlation between simple and unweighted key player measures. The difference in mode size is the absolute difference.

note that the various measures of influence will usually yield divergent results for leading actor identification in networks larger than 50 nodes or with densities above 0.2.

4. Empirical applications

The measures are illustrated using two datasets: the Ndrangheta mafia dataset and the Southern Women dataset (Davis et al., 1941). The former is a two-mode network of mafia members and mafia summits; the latter is a two-mode network of women's attendance at community events. In both sets of empirical demonstrations, I calculate both the bipartite key player measures and the simple key player statistics (as discussed in the Proof of Concept).

4.1 Ndrangheta mafia: fragmentation centrality

The Ndrangheta data describe an Italian mafia network compiled from a large investigation, *Operazione Infinito* (Coutinho, 2016). The data consists of 156 mafia members attending 48 separate summits. Oftentimes, criminologists seek to identify which actors in a criminal network can

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

Table 2. Fragmentation centrality for Ndrangheta mafia data

	Mean	Std. Dev.	Range
Intramodal			
Members (F_i^b)	0.327	0.002	0.320–0.335
Summits (F_q^b)	0.508	0.077	0.485–1.026
Cross-modal			
Members (F_i^c)	0.474	0.003	0.469–0.494
Summits (F_q^c)	0.350	0.218	0.313–1.834
Total			
Members (F_i^*)	0.397	0.002	0.391–0.405
Summits (F_q^*)	0.429	0.148	0.401–1.430

be arrested to maximally disrupt the behaviors of a criminal organization. Fragmentation centrality provides an explicit measure of such vulnerability. In the case of the Ndrangheta data, each measure gives insight into which mafia members or mafia summits should be targeted to maximally disrupt the Ndrangheta criminal organization. The intramodal fragmentation centrality for mafia members gives insight into which members are most vulnerable in the member by member projection, while the intramodal fragmentation centrality for summits gives insight into which summit is most vulnerable in the summit by summit projection. The cross-modal fragmentation centrality for mafia members measures which member, when deleted, will yield the most damage to the summit by summit projection, while the cross-modal fragmentation centrality for summits measures which summit, when deleted, will yield the most damage to the summit by summit projection. The total fragmentation centralities for both members and summits evaluates which summits and which members would yield the most damage to the entire two-mode network when deleted. As in the simulation studies above, subscript i denotes actors in the more populous mode of the network (mafia members), while subscript q denotes actors in the less populous mode (summits).

Table 2 presents summary statistics for the bipartite fragmentation centrality measures. There is little variation in the intramodal fragmentation centrality for mafia members, where the standard deviation of $F_i^b = 0.002$. Alternatively, there is much more variation in the distribution of F_q^b, where the standard deviation is 0.077. While the mean F_q^b (0.508) sits near its minimal value (0.485), the maximum F_q^b is 1.026. The relatively large standard deviation for F_q^b when compared to F_i^b indicates that there is more variation in how deleting a summit would damage the summit by summit projection network than there is how deleting a mafia member would damage the member by member projection. Moreover, the largest F_q^b (1.026) is larger than the largest F_i^b (0.335), indicating that, respective to the size of each projection, removing the leading summit will yield more damage to the summit by summit projected network—as indicated by increases in average distance—than removing the leading mafia member will yield to the member by member projection.

Turning to cross-modal effects, the mean F_i^c is 0.474 with a standard deviation of 0.003 and a range of 0.469–0.494. The standard deviation for F_i^c is slightly larger than the standard deviation for F_i^b (0.002), indicating that there is slightly more variation in how deleting a random mafia member would affect the member by member projection as compared to the summit by summit projection. The cross-modal fragmentation centrality measures for summits also show much variation. The mean F_q^c is 0.350 with a standard deviation of 0.218 and a range of 0.313–1.834, indicating substantial right skew. This result indicates that while removing many of the low-scoring summits will not yield much damage to the mafia by mafia projection, removing the summit with

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

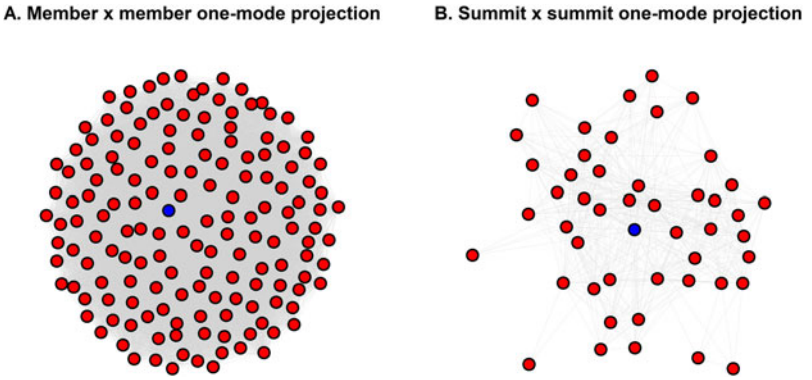


Figure 4. Intramodal fragmentation centrality. Panel A identifies the actor with the maximum F_i^b (0.335) in the member by member projection of the Nradhengta network, while Panel B identifies the actor with the leading F_q^b (1.03). Simple fragmentation centrality statistics were also computed on unweighted variations of the one-mode projections. In the member by member projection, the simple key player statistics were unable to prioritize between the members, yielding a value of 0 for all members. In the summit by summit projection, the simple measures identified the same summit as the weighted metric (maximum simple fragmentation centrality = 0.299).

the largest fragmentation centrality will yield substantial damage. The standard deviation of F_q^c is also roughly three times larger than the standard deviation of F_q^b, indicating that there is more variation in how deleting a random summit would damage the member by member projection than there is in how deleting a random summit would damage the summit by summit projection.

Figure 4 plots the member by member projected network (Panel A) as well as the summit by summit projected network (Panel B). In both panels, the simple and weighted fragmentation centrality measures identified the same actors (blue nodes)—a result which was uncommon in the simulated networks above. This indicates that, in some applied cases, simple and weighted measures of intramodal fragmentation centrality will identify the same actors. Examining the location of actors within the structure of both projected networks yields intuitive insight into each actor's influence: both appear to be centrally located within the projected networks (though this is clearer in Panel B).

One benefit of the bipartite key player measures is that the scores are standardized by the size of the projected networks. Thus, it is possible to compare cross-modal and intramodal fragmentation centrality to evaluate whether, for instance, deleting a high-profile mafia member will yield more or less damage to the member by member projection than deleting a high-profile summit would yield to the member by member projection. A comparison of F_q^c to F_i^b reveals that the smallest F_q^c (0.313) is smaller than the smallest F_i^b (0.320). This indicates that deleting the summit with the smallest cross-modal fragmentation centrality would yield *less* damage to the member by member projection than deleting the mafia member with the smallest intramodal fragmentation centrality, though this difference is small and likely not substantively meaningful in this network. Alternatively, deleting the summit with the largest F_q^c (1.834) will have a larger disruptive effect on the member by member projection than deleting the mafia member with the largest F_i^b (0.335). This indicates that although there are some mafia members who are in more vulnerable locations than some summits, the actor with the most potential to fragment the member by member projection is a summit, rather than a mafia member.

A similar analysis can be conducted for the summit by summit projection. Here, the mafia member with the smallest cross-modal fragmentation centrality F_i^c (0.469) is smaller than the smallest F_q^b (0.485), indicating that deleting the mafia member with the smallest cross-modal fragmentation centrality will yield less damage to the summit by summit projection than deleting the

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

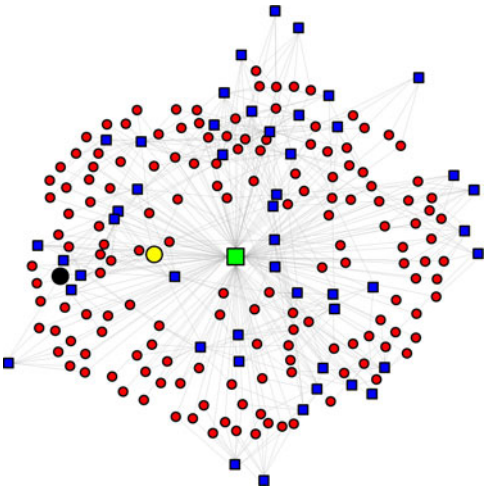


Figure 5. Total and cross-modal fragmentation scores for Ndrangheta mafia network. Blue squares are events; red circles are actors. The yellow circle is the mafia member with the largest F_i^c(0.494) and F_i^* (0.405). The green square is the summit with the largest F_q^* (1.830) and F_q^* (1.430). The black circle is the mafia member with the largest total (0.468) and cross-modal (0.935) simple fragmentation centrality, when calculated from unweighted projections of the bipartite network. The simple calculation of total (0.673) and cross-modal (1.047) fragmentation centrality for summits also identified the green-square as the summit with the maximum fragmentation potential.

summit with the smallest intramodal fragmentation centrality. Likewise, deleting the summit with the largest F_q^b (1.026) will yield more damage to the summit by summit projection than deleting the mafia member with the largest F_i^c (0.494). These results indicate that summits tend to occupy more vulnerable positions with respect to the summit by summit projection than mafia members.

Given that deleting certain summits tends to yield more damage to both the summit by summit projection and the member by member projection, it is likely that the most vulnerable actor in the entire two-mode structure is a summit. This can be assessed by calculating F_i^* and F_q^*. Consistent with the results above, F_q^* tends to be greater in value than F_i^*, reflecting that summits tend to hold greater fragmentation potential (i.e., tend to be more vulnerable) than mafia members: the mean F_q^* is 0.429 with a standard deviation of 0.148 and a range of 0.401–1.430, while the mean F_i^* is 0.397 with a standard deviation of 0.002 and a range of 0.391–0.405. Figure 5 highlights the Ndrangheta mafia members and summits with the leading cross-modal and total fragmentation centralities. Red nodes represent mafia members, while blue squares are summits. The green circle is the summit which is the largest F_q^* and the largest F_q^c. It is centrally located within the network structure, reflecting its high position of influence. This summit was also identified by simple calculations of cross-modal and total fragmentation centrality, further corroborating its high influence. The yellow node is the mafia member with the largest F_i^c (0.494) and F_i^* (0.405). The black node is the mafia member with the largest simple calculations of cross-modal (1.047) and total (0.673) fragmentation centrality. These results reflect that when there are events or actors that wield overwhelming influence (as in the case of summits), the two measures will sometimes identify the same actors. However, when there is less variation among actors, as in the case of the mafia members, the simple and weighted measures will likely not concur and may identify distinct actors. In these instances, the weighted bipartite measures will often be preferable to the simple measures of fragmentation centrality because the weighted measures preserve more information available in the bipartite network structure.

In sum, analyses of the Ndrangheta mafia network illustrate how measures of fragmentation centrality can be used to identify vulnerable locations in a two-mode network. The application illustrates how fragmentation centrality measures can be compared to assess how deleting actors in each respective mode of the network may damage the projections of each mode in the two-mode network. The analyses also show that while simple measures sometimes identify the same actors as weighted measures, they also sometimes diverge. In cases of divergence, it is recommended to use the weighted measures, which preserve more information from the bipartite network when computing one-mode projections.

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

Table 3. Cohesion centrality for southern women data

	Mean	Std. Dev.	Range
Intramodal			
Women (C_i^b)	0.745	0.253	0.326–1.098
Events (C_q^b)	0.919	0.361	0.486–1.612
Cross-modal			
Women (C_i^c)	0.388	0.129	0.175–0.607
Events (C_q^c)	0.299	0.140	0.154–0.592
Total			
Women (C_i^*)	0.567	0.187	0.272–0.842
Events (C_q^*)	0.609	0.246	0.331–1.104

4.2 Southern women: cohesion centrality

The Southern Women data were collected by Davis et al. (1941) in the 1930s. It consists of 18 women who attended 14 social events. The purpose of cohesion centrality is to determine who is the most influential or reachable actor in the network. A signal emanating from this actor (e.g., an idea, a disease) should have the shortest average path in the network structure to all other actors in the intramodal projection, cross-modal projection, and overall bipartite network.

Table 3 shows the cohesion centrality scores for the southern women. Women in the social event network show considerable variation in intramodal cohesion centrality. The average C_i^b is 0.745, with a standard deviation of 0.253 and a range of values stretching between 0.326 and 1.098. This indicates that the woman with the leading C_i^b can disseminate a signal much more quickly to the other women in the women by women projection than the woman with the smallest C_i^b. Turning to the intramodal cohesion centrality of events, the mean C_q^b is 0.919 with a standard deviation of 0.361 and a range of values stretching between 0.486 and 1.612.

As in the case of intramodal cohesion centrality among women, a signal emanating from the event with the leading C_q^b will spread much more quickly than a signal emanating from the event with the smallest C_q^b. In some cases, such as the Davis network, the interpretation of cohesion centrality will have unclear meaning for certain modes of a two-mode network. Events do not diffuse signals without womens' participation. Interpretation of events' cohesion centrality therefore requires reference to the womens' social network. The intramodal cohesion centrality of the event by event projection indicates which event is positioned to most quickly diffuse a signal to the women who are most likely to spread the signal through other events. In less abstract terms, if a disease were to appear at a random event in the Southern Women data, Event 8—the event with the largest intramodal cohesion centrality—would have the greatest potential to spread the disease to the women who attend other events.

Figure 6 demonstrates these measures graphically and compares them to the simple measures of intramodal cohesion centrality. Panel A plots the woman by woman projection. The blue node is Evelyn, who has the largest C_i^b at a value of 1.098. An interesting finding in Panel A is that simple calculations of intramodal cohesion centrality are unable to discriminate between women in the woman by woman projection. While only one woman obtained the highest value of C_i^b, seven women obtained the maximum possible value of 1 when computing simple measures of intramodal cohesion centrality. Six of these women are identified in Panel A as black nodes, while the seventh woman is Evelyn (identified as a blue node because she also maximizes C_i^b). This result indicates that, in some circumstances, simple measures of intramodal cohesion centrality are unable to prioritize actors when examining bipartite networks, while the weighted measures do not encounter this problem.

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

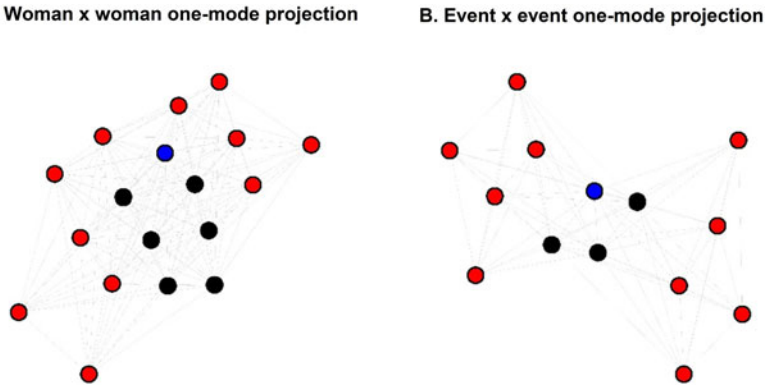


Figure 6. Panel A plots cohesion centrality in the woman by woman one-mode projection of the Davis network, while Panel B plots cohesion centrality in the event by event projection. Blue nodes demark the maximum C_i^b in Panel A (Evelyn, $\max(C_i^b) = 1.098$), and the maximum C_q^b in Panel B (Event 8, $\max(C_q^b) = 1.61$. Black nodes identify those nodes in each projection which obtain the maximum value of 1 when using simple (e.g. Borgatti, 2006) cohesion centrality measures on the one-mode projections in Panels A and B.

This result is replicated in Panel B, which plots the intramodal cohesion centrality of events in the event by event projection. Here, the leading event (Event 8), as measured by C_q^b, is identified as a blue node, with a value of 1.61. This is the event expected to be able to most quickly disseminate a signal to other events in the event by event projection. Simple measures of intramodal cohesion centrality again identify numerous events which obtain the maximum possible value of 1. Here, Event 8 as well as three other events (black nodes) are identified as sharing equal potential to spread a signal through the event by event network. In sum, results in Figure 5 indicate that, unlike weighted measures of cohesion centrality in bipartite networks, simple measures are sometimes unable to prioritize actors when examining intramodal cohesion in two-mode networks.

Turning to cross-modal cohesion centrality, the mean C_i^c is 0.388 with a standard deviation of 0.129 and a range of 0.175–0.607. There is considerable variation in this distribution, indicating that while some women can quickly diffuse a signal throughout the event by event projection, others cannot. As in the case of fragmentation centrality, C_i^c can be compared to C_q^b to evaluate whether signals emanating from women or events more quickly diffuse throughout the event by event projection. The mean C_q^b (0.919) is over twice as large as the mean C_i^c (0.388), indicating that the event with the highest intramodal cohesion centrality can more quickly send a signal throughout the event projection compared to the woman with the highest cross-modal cohesion centrality. However, the smallest C_q^b (0.486) is also smaller than the largest C_i^c (0.607), indicating that some women can diffuse a signal throughout the event by event projection more quickly than some events. Turning to the cross-modal influence of events on the women by women projection, the mean C_q^c is 0.299 (SD = 0.140) with a range of 0.154–0.592. The maximum C_q^c (.592) is smaller than the maximum C_b^c (1.098), indicating that a signal emanating from the most influential event would diffuse less quickly throughout the woman by woman projection than a signal emanating from the women with the highest intramodal centrality. Where C_q^c and C_b^c are Tex notation for the equations.

It is also possible to assess which woman and which event wield the most influence over the entire bipartite network structure. The mean C_i^* is 0.567 with a standard deviation of 0.187 and a range of 0.272–843. The total cohesion centrality scores are comparable for events, with a mean C_q^* of 0.609 (SD = 0.246) and a range of 0.331–1.104. The similar means for the two indicates that both women and events, on average, yield comparable levels of influence over the entire network. Indeed, a *t*-test fails to identify a statistically significant difference in values between the two sets

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

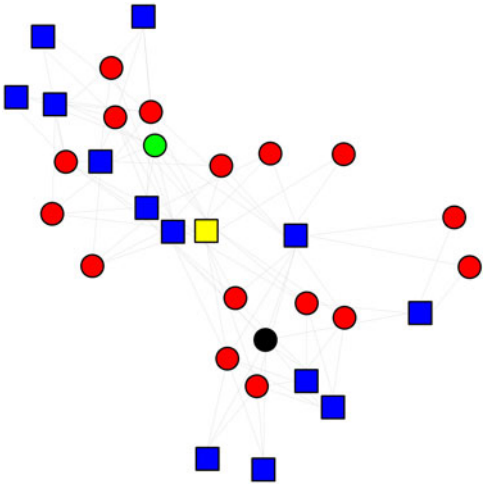


Figure 7. Total cross-modal cohesion for Southern Women Data. Red circles are first mode actors, blue squares are events. The green node is Theresa, an integral woman who maximizes C_i^* (0.842) and C_i^c (0.607), as well as the simple calculation of total cohesion among women (0.692). The black node is Nora, who maximizes the simple calculation of cross-modal cohesion among women. The yellow square is Event 8, which maximizes C_q^* (1.104) and C_q^c (0.592), as well as the simple measures of cross (0.758) and total (0.879) cohesion among events.

of cohesion centralities ($p > 0.1$). That said, the largest total cohesion centrality is for Event 8 ($C_q^* = 1.104$), indicating that the most influential actor across both modes of the Southern Women data is Event 8.

Figure 7 compares simple calculations of cross-modal and total cohesion centrality to weighted bipartite measures. The blue squares are events, while the red nodes are women. The green node is Theresa, the woman who maximizes both C_i^* (0.842) and C_i^c (0.607), as well as the simple calculation of total cohesion among women (0.692). The black node is Nora, a woman who maximizes the simple calculation of cross-modal cohesion among women. The yellow square is Event 8, which maximizes C_q^* (1.104) and C_q^c (0.592), as well as the simple measures of cross (0.758) and total (0.879) cohesion among events. Consistent with the results discussed for fragmentation centrality among summits in the Ndrangheta mafia network, the simple and weighted metrics of cross-modal and total cohesion centrality reach consensus when there is clearly one influential event (Event 8). However, the measures also sometimes identify distinct actors when distances in the projected network are weighted to capture information in the bipartite structure. In the case of cross-modal cohesion, the simple measure identifies a different woman as most influential (Nora) than the weighted measure (Theresa).

In sum, results in empirical applications illustrate how the weighted bipartite key player measures can be used to assess influence in two-mode networks. Key player measures can be used to identify actors which are influential within an intramodal projection, a cross-modal projection, or across the entire bipartite structure. Empirical applications also demonstrate how cross-modal and intramodal key player statistics can be compared to identify whether actors in the first or second mode of a two-mode network yield more or less influence over each projected network. Finally, empirical applications also consider the performance of these measures against simple calculations of key player statistics which do not account for distance in bipartite networks. Results indicate that while the two strategies often identify similar actors, they also often diverge. In some cases, such as intramodal cohesion centrality in the Southern Women projections, simple key player measures were unable to prioritize between sets of four or six vertices, while the weighted measures did not encounter this problem.

5. Discussion

This article proposed a method to identify key players in bipartite networks. The method adapts Borgatti's (2006) cohesion and fragmentation centrality by weighting one-mode projections by mutual affiliations in a two-mode network (Newman, 2001). Geodesic distances in the one-mode

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

projection are determined using a shortest path algorithm, which can be tuned to assign greater or lesser priority to edge weights and the number of intermediaries on a path (Opsahl et al., 2010). With the adjusted geodesic distances, key player statistics can be calculated. Distinct measures were proposed to assess intramodal, total, and cross-modal influence in bipartite networks. The measures are standardized so that centrality scores from actors in different modes of a network or different networks can be compared.

The method contributes to a growing interest in examining two-mode network structure (Everett & Borgatti, 2013; Jasny & Lubell, 2015; Latapy et al., 2008; Opsahl et al., 2010; Opsahl, 2013). Like Latapy et al. (2008) and Everett & Borgatti (2013), it is oriented towards summarizing the three-embedded network structures in affiliation data: $m \times m$, $p \times p$, and $m \times$ p. As such, it offers a useful tool to analyze bipartite networks (e.g Fujimoto et al., 2011; Wang et al., 2009, 2013).

While not discussed at length in this paper, the method can be updated to identifying key player sets. This approach may be useful when an investigator is interested in a subset or clique of actors who yield influence beyond what the linear combination of their key player statistics would suggest. However, the method proposed in this paper may be computationally intensive in large networks, particularly if a researcher is interested in fragmentation centrality. Researchers should use caution and careful theoretical reasoning when specifying the size of the k set to reduce computational burden.

It is also worth discussing here the role of the number of actors in each mode of a bipartite network. Oftentimes, as the absolute difference in the number of actors in each mode of the network increases, the mode with fewer actors will yield more influence. This should not be surprising. In a one-mode network, path lengths will tend to be shorter in smaller networks because there are fewer possible intermediaries between any two actors. Likewise, in a two-mode network, as the number of actors in one mode decreases, the number of shared affiliations between actors in the other mode will also tend to increase, reducing the average distance between two actors in the larger mode. Thus, researchers should bear in mind that when the absolute difference in the sizes of each mode increases, the likelihood that the smaller mode will be more influential increases. Even in these cases, intramodal, cross-modal, and total key player centrality measures in the mode with fewer actors may still be of interest to researchers, depending on the nature of the research question and the goals of key player identification.

When Borgatti (2006) first outlined the key player problem, he proposed examining key players in relation to certain attributes. The method proposed here pushes towards this goal by examining players who are connected through affiliations. In the case of public health, a disease may spread quickly because of an organization which is particularly reachable by both other actors and other organizations. In the case of criminal networks, drug markets may be particularly resilient because of the properties of certain buyers who traffic drugs between sellers. The measures proposed in this article identify such multifaceted positions of influence and vulnerability.

Acknowledgments. I would like to thank David Melamed, Stanley Wasserman, Ron Breiger, and three anonymous reviewers at Network Science for their helpful suggestions and feedback.

Conflict of interest. Scott W. Duxbury has nothing to disclose.

Notes

1 Examples of bipartite networks in each of these contexts are drug markets (buyer–vendor ties), memberships in smoking cessation communities (person–community ties), and interaction of organism with nature reserves (organism–reserve ties).
2 Borgatti (2006) used the terms KPP-Pos (cohesion) and KPP-Neg (fragmentation).
3 Node removal refers to deleting the node and all incident edges from the network.
4 Both datasets were obtained from the UCINET network database.
5 Degree centrality and betweenness centrality have unclear cross-modal analogues, so they are not measured here.

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press

References

Agneessens, F., Borgatti, S. P., & Everett, M. G. (2017). Geodesic based centrality: Unifying the local and the global. *Social Networks*, *49*, 12–26.

An, W., & Liu, Y.H. (2016). keyplayer: An R Package for locating key players in social networks. *The R Journal*, *8*(1), 258–270.

Borgatti, S. P., & Everett, M. G. (1997). Network analysis of 2-mode data. *Social Networks*, *19*(1), 243–269.

Borgatti, S. P. (2006). Identifying sets of key players in a social network. *Computational and Mathematical Organizational Theory*, *12*(1), 21–34.

Breiger, R. L. (1974). The duality of persons and groups. *Social Forces*, *53*(2), 181–190.

Chen, Y., Paul, G., Havlin, S., Liljeros, F., & Stanley, H. E. (2008). Finding a better immunization strategy. *Physical Review Letters*, *101*, 058701.

Cobb, N. K, Graham, A. L., & Abrams, D. B. (2010). Social network structure of a large online community for smoking cessation. *American Journal of Public Health*, *100*(7), 1,282–1,289.

Coutinho, J. (2016). Ndrangheta Mafia 2. *UCINET Data Repository*. Retrieved from https://sites.google.com/site/ucinet-software/datasets/covert-networks/ndranghetamafia2

Davis, A., Gardner, B. B., & Gardney, M. (1941). *Deep South: A social anthropological study of caste and class.* Chicago: University of Chicago Press.

Duijn, P. A. C., Kashirin V., & Sloot, P. M. A. (2014). The relative ineffectiveness of criminal network disruption. *Scientific Reports*, *4*, 4, 238.

Duxbury, S. W, & Haynie, D. L. (2019). Criminal network security: An agent-based approach to evaluating network resilience. *Criminology*, *57*(2), 314–342.

Everett, M. G., & Borgatti, S. P. (2013). The dual-projection approach for two-mode networks. *Social Networks*, *35*(2), 204–210.

Everett, M. G. (2016). Centrality and the dual-projection approach for two-mode social network data. *Methodological Innovations*, *9*(1), 1–8.

Fujimoto, K., Chou, C. P., & Valente, T. W. (2011). The network autocorrelation model using two-mode data: Affiliation exposure and potential bias in the autocorrelation parameter. *Social Networks*, *33*(3), 231–243.

Jasny, L., & Lubell, M. (2015). Two-mode brokerage in policy networks. *Social Networks*, *41*, 36–47.

Larsen, A. G., & Ellersgaard, C. H. (2017). Identifying power elites—k-cores in heterogenous affiliation networks. *Social Networks*, *50*(1), 55–69.

Latapy, M., Magnien, C., & Del Vecchio, N. (2008). Basic notions for the analysis of large two- mode networks. *Social Networks*, *30*(1), 31–48.

Newman, M. E. J. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E*, *64*, 016132.

Newman, M. E. J. (2010). *Networks: An introduction*. Oxford: University of Oxford Publishers.

Opsahl, T., Agneessens, F., & Skvortez, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, *32*(3), 245–251.

Opsahl, T. (2013). Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Social Networks*, *35*(2), 159–167.

Urban, D. L., Minor, E. S., Treml, E. A., & Schick, R. S. (2009). Graph models of habitat mosaics. *Ecology Letters*, *12*(3), 260–273.

Valente, T. W. (2012). Network interventions. *Science*, *337*(6090), 49–53.

Wang, P., Sharpe, K., Robins, G. L., & Pattison, P. (2009). Exponential random graph (p^*) models for affiliation networks. *Social Networks*, *31*(1), 12–25.

Wang, P., Pattison, P., & Robins, G. (2013). Exponential random graph model specifications for bipartite networks—A dependence hierarchy. *Social Networks*, *35*(2), 211–212.

Cite this article: Duxbury S. W. (2020). Identifying key players in bipartite networks. *Network Science* **8**, 42–61. https://doi.org/10.1017/nws.2019.62

https://doi.org/10.1017/nws.2019.62 Published online by Cambridge University Press